REMARKS

Applicant files herewith an amended set of claims, which have been amended to overcome the Examiner's objections.

Claims 10, 12 and 14 have been amended to recite "A data classification apparatus...". Claims 10, 14, 15 and 18 have been amended to ensure that no new matter is added. In these claims, the phrase "an overall strangeness value" has been replaced by the phrase "a single strangeness value", in accordance with page 12, line 10. Claims 15 and 17 now recite "a computer-implemented data classification method" to ensure that they relate to statutory subject matter.

The claims have also been amended to ensure that they are fully distinguished from the prior art, as explained below. In particular, claims 10, 14, 15 and 18 now include the formula by which the single strangeness value is determined by the assay means. This formula is not shown or suggested in any of the prior art documents.

The invention relates to an apparatus for and a method of data classification of unknown items, using a training set of classified examples.

The aim of the invention is not only to predict a classification for an unknown item, but also to provide a measure of confidence in that classification, valid under the iid (independently and identically distributed) assumption. This is the assumption that the training and unknown examples are generated from the same distribution.

As defined in amended claim 10, the data classification apparatus comprises an input device for receiving training classified examples and at least one unclassified example, a memory for storing the examples, an output terminal for outputting a predicted classification for the unclassified example and a processor for identifying the predicted classification, where the processor includes classification allocation means for

allocating potential classifications to each unclassified example and for generating a plurality of classification sets, each set containing the plurality of training classified examples (l) with the classification and at least one unclassified example (l+1) with its allocated potential classification, an assay means including an example valuation device which determines individual strangeness values (α_i) for each training classified example (i=1,2...l) and at least one unclassified example (i=l+1) having an allocated potential classification (y), the assay means determining a single strangeness value (d(y)) valid under the iid assumption for each classification set in dependence on the individual strangeness values (α_i) of each example by a given formula, a comparative device for selecting the classification set to which the most likely potential classification for the unclassified example belongs on the basis of the single strangeness value assigned by the assay means, and a strength of prediction monitoring device for determining a confidence value for the predicted classification on the basis of the single strangeness value assigned to the classification set to which the second most likely potential classification of the unclassified example belongs.

The formula $d(y) = \frac{|\{i : \alpha_i \geq \alpha_{l+1}\}|}{l+1}$, where i=1,2....l, l+1, calculates the single strangeness value by counting how many of the individual strangeness values of the training examples are greater than or equal to the individual strangeness value of the unclassified example, and dividing this by the total number of examples.

Thus, the apparatus starts with training classified examples, and uses these to predict the classification of an unclassified example. The classification allocation mean allocates potential classifications to each unclassified example, and generates a number of classification sets, each containing the said training examples with their classification, and at least one unclassified example with a potential classification. Each pair of the example and classification in the set is tested by the assay means to determine its individual strangeness value. That is a measure of how strange it would be if the classification were correct. The assay means then determines a single strangeness value,

valid under the iid assumption, for each classification set, in dependence on the individual strangeness values. The comparative device compares the single strangeness value for all the sets, to see which is most likely to be correct for the unclassified example. The strength of prediction monitoring device then looks at the second most likely classification, on the basis of the single strangeness values, to give a measure of confidence that the most likely classification is in fact correct. Thus, the greater the difference between the single strangeness values of the most likely potential classification and the second most likely potential classification, the greater is the measure of confidence that the most likely classification is correct.

We submit that this combination of features, and in particular the formula, is not shown in any of the prior art documents.

Applicant agrees that Mizuno shows an input device, a memory, an output terminal and a processor, with the processor including classification allocation means and assay means. However, applicant submits that Mizuno calculates an error between a classification assigned to an unknown pattern and a correct classification, and initiates re-training when the average error exceeds a predetermined value. This copes with the possibility that the unknown patterns are generated by a different mechanism from the training patterns (see column 1, lines 9 to 19). This differs from the invention as claimed, which specifies that the single strangeness value is valid under the iid assumption, in that the unknown and training examples are generated from the same distribution. It is submitted that column 7, lines 1 to 12 of Mizuno do not show the determination of an individual strangeness value for each training example and an unknown example and then the determination of a single strangeness value for each classification set, in dependence on the individual strangeness values. In particular, Mizuno does not show the formula now included in amended claim 10. Applicant agrees that Mizuno does not teach a strength of prediction monitoring device for determining a confidence value for the most likely classification, on the basis of the single strangeness value of the second most likely potential classification.

It is submitted therefore that claim 10 is clearly novel over Mizuno. Further, Mizuno does not show or suggest determining a single strangeness value valid under the iid assumption, and indeed is concerned with the possibility that the unknown patterns are generated from a different distribution, rather than the same distribution. Thus, Mizuno teaches away from the invention, rather than towards it. In addition, Mizuno does not show or suggest determining individual strangeness values and then a single strangeness value, either in general or on the basis of the claimed formula, nor the strength of prediction monitoring device for a confidence value on the basis of the single strangeness value of the second most likely potential classification.

Tsuboka relates to pattern recognition employing the Hidden Markov Model, and is particularly relevant to speech recognition or other time series. In such pattern recognition it is necessary to identify a spoken word, and to decide which of a number of classified words it is most likely to be. Examples of classified words are put into clusters and then the probability is calculated that the word belongs to a particular cluster. Tsuboka recognises that a pattern to be recognised has different probabilities of belonging to different clusters, and calculates the probability in various ways, using a comparative device to select the most likely classification. However, the aim of Tsuboka is simply to find the cluster with the highest probability, for accuracy of recognition.

Applicant submits that firstly Tsuboka does not use the iid assumption, as claimed by the applicant. Secondly, applicant submits that Tsuboka does not use the second most likely classification as a strength of prediction monitoring device. The passage indicated by the Examiner (column 11 lines 11 to 20) appears simply to say that a pattern could be in more than one classification with different probabilities. The subsequent portion of the specification indicates how the most likely classification is then computed. Tsuboka clearly does not show the use of the formula, as now included in amended claim 10.

Thus, it is submitted that claim 10 is clearly novel over Tsuboka. Further, Tsuboka does not show or suggest using the second most likely allocated potential

classification as a strength of prediction monitoring device for determining a confidence value of the predicted classification. Claim 10 is therefore clearly inventive in view of Mizuno.

With regards to a combination of Mizuno and Tsuboka, it is submitted that they would not lead to the present invention. Firstly, neither is concerned with ensuring that the strangeness value is valid under the iid assumption, and indeed Mizuno is concerned with the possibility that the unknown patterns are generated from a different distribution. In addition, neither Mizuno nor Tsuboka shows the use of the second most likely classification as the strength of prediction monitoring device to provide a confidence value for the correctness of the predicted classification.

Thus, a combination of Mizuno and Tsuboka, even though they are both concerned with maintaining accuracy of the classification, simply would not lead to the present invention, as claimed in claim 10, because of the lack of provision of a confidence value using the second most likely classification, and the fact that the strangeness values are not valid under the iid assumption. The combination also fails to show the use of the formula enumerated in the claim.

Applicant submits therefore that claim 10 is clearly distinguished from all the prior art, and is therefore allowable. Claims 14, 15 and 18 have the same limitations as claim 10, and it is therefore submitted that these are also allowable.

As claims 12 and 17 are dependent on claims 10 and 15, it is submitted that these are allowable to the same extent.

In re Appln. of:Gammerman et al.
Application No. 09/831,262

Conclusion

The application is considered in good and proper form for allowance, and the Examiner is respectfully requested to pass this application to issue. If, in the opinion of the Examiner, a telephone conference would expedite the prosecution of the subject application, the Examiner is invited to call the undersigned attorney.

Respectfully submitted,

Phillip M. Pippenger, Reg. No. 46,055
LEYDIG, VOIT & MAYER, LTD.
Two Prudential Plaza, Suite 4900
180 North Stetson Avenue
Chicago, Illinois 60601-6780
(312) 616-5600 (telephone)
(312) 616-5700 (facsimile)

Date: November 18, 2004

Amendment or ROA - Final (Revised 10/21/2004)